DOCUMENT RESUME

ED 167 627 TM 008 482

AUTHOR Nagy, Philip
TITLE Construct Validity of Cognitive Structures: A Comparison of Multidimensional Methods.
PUB DATE Mar 78
NOTE 25p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978) ; Best copy available

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Cognitive Development; *Cognitive Measurement; Cognitive Style; Complexity Level; Data Analysis; Foreign Countries; *Goodness of Fit; *Multidimensional Scaling; Personality Assessment; *Scientific Concepts; Senior High Schools; *Validity
IDENTIFIERS Canada; *Construct Validity

ABSTRACT

This study assessed the construct validity of a cognitive structure interpretation of multidimensional scaling solutions of concept similarity data. Using high school subjects, convergent validity was assessed through correspondence of scaling solutions of three similarity rating tasks; word association, similarity judgment, and semantic differential. Further evidence was sought by grouping subjects within each data gathering technique by similarity of response using, separately, Tucker's and Carroll's individual scaling techniques. Subgroup membership comparisons were made within data gathering techniques across scaling techniques, and within scaling techniques across data gathering techniques. Characterization of subgroup members was attempted through several variables, including developmental level, field independence, cognitive complexity, and measures of ability to cope with the similarity tasks. On full group data, fit was excellent between word association and similarity judgment data, and poorer with semantic differential data. Judging by the fit among the interpretability of subgroup solutions, both individual scaling techniques formed groups successfully. There was considerable communality of subgroup membership across similarity tasks in the Carroll analysis only. Characterization of subgroup members by the chosen variables was unsuccessful. Due to error-full data, only Carroll analysis and not Kruskal analysis produced interpretable dimensions on the subgroup data. (Author/RD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED167627

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

CONSTRUCT VALIDITY OF COGNITIVE STRUCTURES: A COMPARISON OF MULTIDIMENSIONAL METHODS

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Philip Nagy

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Philip Nagy
Newfoundland Dept. of Education
St. John's, Nfld.

TM008 442

Presented at the Annual Conference, American Educational Research Association. Toronto, Ontario. March, 1978. Session 27.12.

ABSTRACT

This study assessed the construct validity of a cognitive structure interpretation of multidimensional scaling solutions of concept similarity data. Using high school subjects, convergent validity was assessed through correspondence of scaling solutions of three similarity rating tasks: word association, similarity judgment, and semantic differential.

Further evidence was sought by grouping subjects within each data gathering technique by similarity of response using, separately, Tucker's and Carroll's individual scaling techniques. Subgroup membership comparisons were made within data gathering techniques across scaling techniques, and within scaling techniques across data gathering techniques.

Characterization of subgroup members was attempted through several variables, including developmental level, field independence, cognitive complexity, and measures of ability to cope with the similarity tasks.

On full group data, fit was excellent between word association and similarity judgment data, and poorer with semantic differential data. Judging by the fit among and interpretability of subgroup solutions, both individual scaling techniques formed groups successfully. There was considerable commonality of subgroup membership across similarity tasks in the Carroll analysis only. Characterization of subgroup members by the chosen variables was unsuccassful. Due to error-full data, Carroll analysis only and not Kruskal analysis produced interpretable dimensions on the subgroup data.

Construct Validity of Cognitive Structures:
a Comparison of Multidimensional Methods

INTRODUCTION

The purpose of this investigation was to assess the validity of three methods of uncovering "cognitive structures" from perceived interrelationships among a set of science concepts. The validation follows, in spirit, that of Campbell and Fiske (1959). Convergent validity was assessed by analysis of interrelationships produced by the same subjects on different similarity rating tasks, and by analysis of those produced by the same subjects on the same similarity rating task but analysed by different scaling methods. Discriminant validity was assessed (only indirectly) through analysis of relationships produced by different groups of subjects. These groups were formed using the individualized scaling models of Tucker and Messick (1963) and Carroll and Chang (1970). Characterization of the members of these groups was attempted through a set of personality measures, some independent of and others dependent upon the scaling techniques.

The construct validation may be viewed as a two-step process. The first step validates the methods of gathering similarity ratings, through the establishment of the degree of correspondence of results across the different similarity rating tasks. The second step is the validation of the recovered scaling solutions, attempting to establish them as representing "cognitive structures", in that they represent some psychological reality for the subject. This second step was attacked by two methods. First, scaling solutions were analysed in terms of the meanings of the example concepts on which similarity data was gathered. Second, the personality measures of groups of subjects who perceived the example concepts from different perspectives were analysed in an attempt to link personality characteristics with these perspectives.

PROCEDURE

Sample

The sample was a group of 78 grade nine and 54 grade twelve students in four schools in Edmonton, Alberta. The sample originally chosen was considerably larger, but since subjects were retained in the sample only if they completed six tests administered on three separate occasions, there was a large (50%) attrition rate. Since there was no attempt in this study to generalize perceptions of science concepts to "typical" classes, for the purposes of investigating the data gathering and scaling techniques only, the attrition rate was deemed acceptable.

Instrumentation

The study used as example concepts a set of 15 syntactical (as opposed to substantive) concepts within the domain of the scientific method. These were:

conclusion	discovery	evidence
experiment	explanation	fact
hypothesis	imagination	interpretation
investigation	law	proof
puzzle	question	theory

Four matrices of similarity ratings among these 15 concepts were produced from the responses of each subject as follows:

1) A constrained word association test was administered in which subjects were asked to respond continually, in writing for one minute per concept, to each of the 15 concepts, the constraint being that they were to "respond while thinking of the scientific meaning of the stimulus word." Using the Garskof and Houston (1963) method with small modification, relatedness indices with values between 0.00 and 1.00 were calculated between all pairs of words based upon the overlap of the lists of responses to each concept. Since the stimulus word itself is included in the list, overlap can never be perfect as the first two words can never be the same on both lists. A correction factor was used

by Garskof and Houston to allow a theoretical maximum of 1.00 if each word elicits the other, followed by identical lists of responses. This correction factor has the disadvantage of producing the same index, 1.00, under the described circumstances whether the list of identical responses is one or ten words long. Since it was felt that production of a longer list of identical responses should indicate a higher proximity of concepts in "concept space", the correction factor was not used in this study. Thus, longer identical lists, had there been any (the highest relatedness index produced by any subject was about 0.70), would have produced slightly higher relatedness indices. The correction factor would have changed a typical index from, for example, 0.439 to 0.444.

2) A second matrix of similarity data was also produced from the association responses. There were very large between individual differences in both fluency of response and average relatedness index. The average relatedness index may be taken as an inverse indicator of the size of the concept pool from which the subject responded. Each subject's matrix was renormed to a mean of 10.00 and a standard deviation of 2.00 to eliminate differences in fluency and average relatedness index. Without such renorming, individuals with the highest mean and standard deviation of relatedness indices would have dominated any group average calculations. However, since this fluency factor may have been theoretically important, data was analysed both scaled (described here) and unscaled (described in (1) above).

3) A similarity judgment task (Torgerson, 1958) was administered, in which subjects were asked to rate 105 possible pairs of concepts by making a mark on a line labelled "closely related-unrelated". Measurements from one end of the line produced a set of similarity ratings directly. These values were also rescaled to a mean of 10.00 and a standard deviation of 2.00 to remove patterns of response bias. Since such patterns were not considered theore-

tically important, the data was analysed only in the rescaled form.

4) A semantic differential scale (Osgood et al, 1957) consisting of twelve pairs of descriptors was used to rate each of the fifteen concepts. The results were analysed by first producing an intercorrelation matrix for the ratings of the concepts based upon the patterns of responses produced, and then using these intercorrelations as measures of similarity. Again, to eliminate theoretically unimportant response patterns, the matrices were renormed to a mean of 10.00 and a standard deviation of 2.00.

An attempt was made to characterize subjects with a variety of personality measures of potential theoretical importance. Three tests were used for this purpose:

1) A test of developmental level was compiled from existing items, with some modifications (Karplus and Lavatelli, 1969; Hobbs, 1975). Time constraints limited the test to four items, which were (a) conservation of volume; (b) combination of variables; (c) and (d) two controlling variables tasks;

2) A published test of field independence, the Hidden Figures Test (French et al, 1965) was administered;

3) A test of cognitive complexity was designed, following the model of Seaman and Koenig (1974), but limited to a six by six matrix.

The above variables were deemed to be of theoretical importance to the study, in that the three variables might reasonably be expected to play a role in explaining a person's method of relating concepts together in a "cognitive structure". During the course of the investigation, other variables were generated from the data which might have had explanatory power. For example, subjects varied in ability to cope with the word association. Because of this, the following variables were generated: number of responses, number and percent of responses which were original keywords, and personal consistency scores on each of the three similarity tasks (to be described below).

Traditional variables of sex, science grade from the previous year, and I.Q. (grade nine only) were also investigated.

Design

Each of the concept relationship tasks was paired, on the basis of length, for administration with one of the personality tests. The pairings were: Hidden Figures and similarity judgment, developmental level and word association, and cognitive complexity and semantic differential. The personality test was always administered first.

With three testing sessions needed per group, it was planned to test six classes at each grade, each class receiving the sessions in one of the six possible orders. Due to difficulty in finding grade twelve teachers willing to sacrifice three days to research in April of the graduating year, the original design could be followed in grade nine only. Only five classes were found for grade twelve, with one of the six orders of administration being discarded.

Analytic Techniques

In each grade, four group average similarity rating matrices were formed on the basis of the four data sets. For brevity, these eight sets will be referred to as: constrained association scaled (CAS9 and CAS12) and unscaled (CAU9 and CAU12), similarity judgment (SJ9 and SJ12), and semantic differential (SD9 and SD12). An attempt was made to establish dimensionality of the solutions using the Kruskal scaling technique (1964a, 1964b) on each of the group average matrices. The "elbow" criterion was not useful, as would be expected on the basis of the behaviour of the stress function under error conditions (Wagenaar and Padmos, 1971; Isaac and Poor, 1974). On the basis of interpretability as much as stress, all results are reported in three dimensions. Reasons for this will be discussed below.

Comparisons between scaling solutions was done by the nonmetric

orthogonal Procrustes goodness-of-fit procedure (Schonemann and Carroll, 1970; Lingoes and Schonemann, 1974). Rather than the measure of fit used in this technique, the more familiar Pearson correlation derived from it was reported.

Each of the eight sets of individual matrices was analysed by both the Tucker technique and the Carroll. These techniques group subjects according to similarity of response matrices. Subjects were grouped by their loadings in "person space" within each grade. These subgroups were then examined for similarity within data gathering methods and across scaling methods, and vice versa. Matrices within each subgroup were averaged, with the result analysed by the Kruskal technique. As explained below, subgroups were also analysed by the Carroll method. Attempts were made to interpret dimensions in terms of concept meanings, and to characterize members of the subgroups according to personality measures obtained.

RESULTS AND DISCUSSION

Reliability of Instruments

The reliabilities of the instruments used were estimated in a variety of ways, some with recognized limitations. A separate group of 18 grade nine students provided test-retest (one month) data on the constrained association and cognitive complexity instruments. Repeated pages in the test booklets gave further correlational evidence of reliability of the similarity judgment and semantic differential instruments, while a similar technique in the constrained association test gave an estimate in terms of the Garskof and Houston index. These three estimates are reported in Table 1 as personal consistencies. Reliability for the developmental level test was estimated by KR-20, and the Hidden Figures test by split halves. The cognitive complexity test was dropped due to low reliability. Reliabilities are summarized in Table 1.

****
Table 1 about here
****

Group Average Results

Kruskal scaling was performed on the group average matrices. The stress 1 values for two to six dimensional solutions are reported in Table 2. Although the "elbow" criterion is not entirely appropriate, it would seem to indicate that CAS9, SJ9, and SD12 require three dimensions, CAU12, CAS12, and SD9 four, and offer no clear decision for the remaining two cases.

****
Table 2 about here
****

A decision was made to report only three dimensional solutions for the following reasons. First, according to Isaac and Poor (1974), the stress values for all eight cases would be acceptable for a 40% error situation, which is reasonable in light of the reliabilities. Second, based on pilot work, three dimensions would be easier to interpret than four. Third, since comparisons between subgroups are to be made, subgroups might be expected to require fewer dimensions due to their homogeneity. Fourth (in retrospect), Carroll analysis showed no significant improvement in correlation with the original data in moving from three to four dimensions.

Due to the bulky nature of the data in this study, comprising over 100 pages of tables, only example results are reported here. Results given are representative of those omitted. Complete results are available from the author on request.

Table 3 contains sample scaling solutions for the data. Since dimensions in Kruskal analysis are arbitrary, all solutions are reported rotated to a common target, arbitrarily chosen as the SJ9 results.

****
Table 3 about here
****

Although it was not a focus of this study, attempts were made to interpret dimensions of these rotated solutions. The labels chosen were "certainty", "creativity", "theoretical-practical", and "temporal". Subtle and arbitrary distinctions were made in order to create distinguishable labels. For example, high loading of the concept "conclusion" indicated a "temporal" rather than a "certainty" dimension. When "theory" or "law" was opposed to "conclusion" or "fact", a "theoretical-practical" dimension was named. A summary of such interpretations is given in Table 4. In general, the first two dimensions of all solutions are labelled the same. This consistency is interpreted as convergent validity evidence. It should be noted that the semantic differential solutions required the most stretching of the dimension naming conventions.

****
Table 4 about here
****

Table 5 reports further evidence of convergent validity in the form of goodness-of-fit measures of the scaling results. The table also reports on the effect of scaling the association matrices for fluency differences, and the difference between the two grades. In general, the association data and similarity judgment data fit each other better than either fits the semantic differential data. Between grade differences may be interpreted as discriminant validity, but the numerical evidence is weak, even granted that there is no estimate of how much difference three years exposure to science should make.

****
Table 5 about here
****

Tucker Scaling Results

Table 6 presents the unrotated factor loadings for the first six factors for each data set. Since the first factor is a mean factor in data sets 3 to 8, due to the initial norming, percentage loadings of all but the

first factor are expressed as percentages of variance left after the first factor is taken out. This explains percentages totalling more than 100%.

****
Table 6 about here
****

Since each factor retained would form a group, and scree tests were of no help, it was arbitrarily decided to retain four factors (factors 2 to 5 for data sets 3 to 8) for varimax rotation. After rotation, subjects were grouped according to their highest loading, with the following exceptions: those loading greater than 1.00 on more than one factor; those loading less than 0.50 on all factors; and those whose highest loading was less than 1.50 times their next highest were all eliminated. In this manner, an average of 61% of the subjects were placed in four groups within each data set in a mutually exclusive but not exhaustive manner. As there were no commonalities of group membership, these memberships are not reproduced here. Had there been commonality, this would have presented compelling evidence for convergent validity of the scaling solutions

Table 7 gives stress values for the Kruskal analyses of these subgroup average matrices. Since those retained in groups would be expected to have produced more homogeneous and less error-prone data than those omitted, stress values should have been lower than in Table 2. That this was not so must be due to higher error in the average data of the smaller group, as will be discussed below.

****
Table 7 about here
****

Table 8 reports goodness-of-fit correlations among a full group solution and the four corresponding subgroup solutions. Although the Tucker technique has identified subgroups with different perceptions, a search of the 96 dimensions produced found only the occasional interpretable dimension.

This point will also be returned to.

****
Table 8 about here
****

Due to difficulties in matrix inversion due to near redundancy of variables, the personality measures could not be investigated by the appropriate technique, multivariate analysis of variance. As an alternative, one-way Anova at the 0.01 level was chosen. Nine of the 116 analyses of variance were significant, and are reported in Table 9. As can be seen, only the internally generated variables produced significant results. It is of interest that fluency formed a basis for grouping in the unscaled association data which did not persist in the scaled data, despite the high correlation between the overall group results in the two cases. The rescaling, therefore, did prevent the formation of a group on the basis of fluency.

****
Table 9 about here
****

There are several possible reasons for the lack of interpretability of the subgroup solutions. First, error in the data may have confounded the results. Or, if perceptions were constrained by the "right answers", subjects close to the group average may have loaded on several dimensions, and been eliminated by the grouping criteria. Thus, those grouped might be those who deviate substantially from the group average, and who might have the most error-prone data. Since the Kruskal technique seems unable to operate in the presence of high error levels, perhaps the Wagenaar and Padmos criteria for acceptable stress are too lenient.

If the Tucker technique has failed, then the personality results are to be expected. If not however, then we may have chosen the wrong variables, measured them inaccurately, or they may be irrelevant to concept perception. The Carroll scaling results will shed light on these possibilities.

Carroll Scaling Results

Carroll scaling on the full group data sets was done both from a random number start and a Kruskal group solution start. Since correlation was 0.99 between the two, only the former is reported. Table 10 reports the fit of the Carroll group average result to the Kruskal result and to the original data. Perfect fit was not expected between the Kruskal and Carroll results, as the former weights all subjects equally, while the latter gives more weight to those closer to the group average.

****
Table 10 about here
****

In support of the reporting of results in three dimensions rather than four, average correlation of Carroll results with data, 0.58 in three dimensions, rose to only 0.62 in four.

Rather than producing subgroup solutions by weighted composites of the Carroll loadings, analysis proceeded in a manner analogous to the Tucker method. Following subgroup formation, the subgroup average matrices were analysed by the Kruskal technique. Since the subject dimensions in Carroll analysis directly reflect the importance of corresponding dimensions in concept space, the group formation criteria differed from the Tucker analysis. For grade nine, the ten highest of 78 (grade twelve, eight of 54) loadings were used to form each of the three groups. This procedure placed 41% of subjects in three groups, which is of the same order of magnitude as the 61% in four groups of the Tucker analysis. The Carroll groups were not mutually exclusive.

Since there is substantial commonality, group membership is summarized in Table 11. With the exception of two pairs of subjects, grouping is the same in all subgroups of CAU9 and CAS9, and in CAU12 and CAS12. CAU9 group one and SJ9 group three have six members in common. Overlap can also be seen

between SJ9 group two and both CAU9 and CAS9 groups one and two, SD9 group one and SD9 group three, CAU12 group one and SJ12 group one, and CAS12 group three and SJ12 group two. In contrast to the Tucker results, these results offer strong evidence both for the construct validity of the cognitive structures uncovered, and for the power of the Carroll technique in error-full data.

****
Table 11 about here
****

The subgroup average matrices produced were analysed by the Kruskal technique. Stress values were high, similar to those for the Tucker analyses. Measures of fit among the subgroup solutions and with the corresponding full group are reported in Table 12. Average fit was 0.48, slightly better than the 0.44 in the Tucker case.

****
Table 12 about here
****

Attempts to label dimensions were largely fruitless, as with the Tucker results. The personality measures again had no explanatory power. Despite the temptation to conclude that neither individual technique was operating as desired under these circumstances, one bit of evidence prompted further investigation. Those subjects included in subgroups by the Carroll scaling had a higher correlation of the solution with their original data, 0.68, than did the full group, at 0.58. That is, the Carroll technique was successfully isolating those nearer the group average, and was not operating on random error.

This fact prompted further investigation of both the Tucker subgroup data and the Carroll subgroup data using the Carroll technique in place of the Kruskal technique. As can be seen from Table 13, about 75% of the isolated dimensions can be assigned meaning on the basis of the criteria used in attempting a similar naming for the corresponding Kruskal solutions. In

Table 14, correlations between the Carroll solutions and the corresponding Kruskal solutions show that the two results bear little resemblance to one another. On this evidence, it seems clear that the Tucker and Carroll techniques are grouping on some real but unidentified individual differences.

****
Tables 13 and 14 about here
****

The goodness-of-fit results among the Carroll solutions of the Tucker subgroups and corresponding full group are reported in Table 15, along with the same results for the Carroll groups. There is a reasonable balance between no individual differences ($r = 1.00$) and differences so extreme (as in Tables 8 and 12) that the results must be suspect. For the purposes of addressing the hypotheses of this study, the possibility of labelling the dimensions of the subgroup solutions, and the isolation of real differences in concept perception offer evidence for the validity of a cognitive structure interpretation of individual scaling solutions. Commonality of subgroup membership across data gathering techniques within the Carroll analysis, but not within the Tucker analysis, is evidence that the Carroll technique is more robust with respect to error than is the Tucker technique.

****
Table 15 about here
****

SUMMARY

The purpose of this study was to seek convergent validity of a cognitive structure interpretation of multidimensional scaling solutions of concept similarity data. Three data acquisition techniques and two analytic techniques for subject grouping were used. Various personality measures were used unsuccessfully to characterize members of subgroups. There was commonality of subgroup membership within the Carroll analysis, but not within the Tucker analysis. On subgroup data, the Kruskal technique gave uninterpretable loadings in concept space. Similar analysis by the Carroll technique, however,

gave interpretable dimensions. On the group average data, convergence of concept similarity data was good. This convergence was taken as evidence for the cognitive structure interpretation of multidimensional scaling solutions on concept similarity data.

REFERENCES

Campbell, D.T. and Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Carroll, J.D. and Chang, J.-J. An analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. Psychometrika, 1970, 35, 283-319.

French, J R., Ekstorm, R.N., and Price, L. Kit of Reference Tests for Cognitive Factors. Princeton, N.J.: Educational Testing Service, 1965.

Garskof, B.E. and Houston, J.P. Measurement of verbal relatedness: an idiographic approach. Psychological Review, 1963, 70, 277-288.

Hobbs, Edward D. Formal operations in secondary students: a test of the idea of intellectual structural limitations. Unpublished Ph.D. dissertation. University of Alberta, 1975.

Isaac, P.D. and Poor, D.S. On the determination of appropriate dimensionality in data with error. Psychometrika, 1974, 39, 91-109.

Karplus, R. and Lavetelli, C. The developmental theory of Piaget: Conservation. San Francisco: John Davidson Film Producers, 1969.

Kruskal, J.B. Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis. Psychometrika, 1964a, 29, 1-27.

Kruskal, J.B. Nonmetric multidimensional scaling: a numerical method. Psychometrika, 1964b, 29, 115-129.

Lingoes, J.C. and Schonemann, P.H. Alternative measures of fit for the Schonemann-Carroll matrix fitting algorithm. Psychometrika, 1974, 39, 423-427.

Osgood, Charles E., Suci, G.J , and Tannenbaum, P.H. The Measurement of Meaning. Urbana, Ill.: University of Illinois Press, 1957.

Schonemann, P.H. and Carroll, R.M. Fitting one matrix to another under choice of a central dilation and a rigid rotation. Psychometrika, 1970, 35, 245-255.

Seaman, J.M. and Koenig, F. A comparison of measures of cognitive complexity. Sociometry, 1974, 37, 375-390.

Torgerson, W.S. Theory and Methods of Scaling. NewYork: Wiley, 1958.

Tucker, L.R. and Messick, S. An individual differences model for multidimensional scaling. Psychometrika, 1963, 28, 333-367.

Wagenaar, W.A. and Padmos, P. Quantitative interpretation of stress in Kruskal's multidimensional scaling technique. British Journal of Mathematical and Statistical Psychology, 1971, 34, 101-110.

TABLE 1

SUMMARY OF RELIABILITY ESTIMATES

Test	Method	Grade [illegible]	Grade 12
CAU	Test-retest (data)	0.76	
CAS	Test-retest (data)	0.76	
CAU	Test-retest (soln)	0.70	
CAS	Test-retest (soln)	0.53	
CAU	Personal consistency	0.46	0.42
SJ	Personal consistency	0.56	0.51
SD	Personal consistency	0.60	0.64
Dev. Lev.	KR-20	0.45	0.50
Hid. Fig.	Split halves	0.77	0.78
Cog. Com.	Test-retest	-0.11	

TABLE 2

STRESS 1 VALUES FOR FULL GROUP AVERAGE KRUSKAL SOLUTIONS

	Dimension				
	6	5	4	3	2
Data Set					
1. CAU9	4.8	5.0	4.9	8.1	11.4
2. CAU12	4.9	5.0	5.6	10.2	13.4
3. CAS9	4.8	5.0	5.0	6.2	14.0
4. CAS12	5.0	5.0	6.1	10.8	15.8
5. SJ9	5.0	5.0	5.0	6.1	14.1
6. SJ12	4.9	4.9	5.1	7.8	15.5
7. SD9	4 7	4.8	5.0	12.7	16.7
8. SD12	4.9	4.9	4.9	6.2	11.9

TABLE 3

EXAMPLE KRUSKAL SOLUTIONS OF FULL GROUP
DATA ROTATED TO SAME TARGET

	Dimensions					
	1	2	3	1	2	3
Concepts	CAS9 data			SJ12 data		
Conc	-0.07	0.07	-0.38	-0.19	-0.18	0.05
Disc	0.10	-0.26	-0.40	0.10	-0.04	-0.23
Evid	-0.23	-0.21	-0.05	-0.17	-0.15	-0.18
Expt	0.19	-0.10	-0.05	0.21	-0.08	-0.01
Expl	-0.09	0.12	-0.11	-0.17	0.04	-0.11
Fact	-0.32	-0.18	-0.04	-0.28	-0.22	-0.24
Hypo	0.12	0.27	0.15	0.07	0.18	0.28
Imag	0.07	0.22	0.56	0.24	0.54	0.27
Inte	-0.03	0.35	-0.07	-0.15	0.27	-0.07
Inve	0.20	-0.16	-0.03	0.17	-0.16	0.03
Law	-0.40	-0.34	0.19	-0.45	-0.20	-0.09
Proo	-0.23	-0.22	-0.06	-0.20	-0.25	-0.15
Puzz	0.50	0.15	0.16	0.55	0.09	0.14
Ques	0.28	0.16	-0.05	0.35	0.09	0.04
Theo	-0.09	0.12	0.19	-0.09	0.08	0.25

TABLE 4

SUMMARY OF DIMENSION INTERPRETATIONS FOR KRUSKAL
SOLUTIONS OF FULL GROUP DATA

Data Set	Dimension		
	1	2	3
CAU9	1	2	4
CAU12	1	2	0
CAS9	1	2	3
CAS12	1	2	0
SJ9	1	2	1
SJ12	1	2	3
SD9	1	2	0
SD12	1	3	0

0 - No clear interpretation; 1 - Certainty; 2 - Creativity;
3 - Theoretical-practical; 4 - Temporal

TABLE 5

GOODNESS-OF-FIT MEASURES FOR KRUSKAL SOLUTIONS
OF FULL GROUP AVERAGE DATA

Rotated Matrix	Target Matrix	r
Convergent Validity		
CAS9	SJ9	0.91
CAS9	SD9	0.62
SJ9	SD9	0.64
CAS12	SJ12	0.89
CAS12	SD12	0.76
SJ12	SD12	0.76
Effect of Scaling		
CAU9	CAS9	0.94
CAU12	CAS12	0.97
Grade Differences		
CAS9	CAS12	0.78
SJ9	SJ12	0.92
SD9	SD12	0.70

TABLE 6

UNROTATED TUCKER FACTORING-PERCENTAGE TOTAL VARIANCE
ACCOUNTED FOR BY FIRST SIX FACTORS
FOR EACH DATA SET SEPARATELY

	Factors					
	1	2	3	4	5	6
Data Set						
CAU9	55.5	5.8	3.1	2.8	2.5	2.4
CAU12	62.7	5.1	3.3	2.4	2.3	2.2
CAS9	97.2	8.0	7.4	6.2	5.3	4.8
CAS12	97.1	8.2	6.6	6.5	5.9	5.8
SJ9	97.3	7.9	6.1	5.2	4.4	4.0
SJ12	97.2	7.6	5.5	5.4	5.0	4.5
SD9	96.5	7.8	6.3	5.6	5.3	4.5
SD12	96.6	9.3	8.8	7.4	5.7	5.5

For data sets 3-8, percentages calculated excluding first factor.

TABLE 7
STRESS VALUES FOR KRUSKAL SCALING OF
TUCKER SUBGROUP AVERAGE MATRICES
(EXAMPLE - SJ9)

Group	Dimension 6	5	4	3	2
1	10.7	13.5	16.5	21.9	31.0
2	11.3	12.5	15.9	21.4	29.6
3	10.8	13.2	17.2	22.5	30.1
4	10.5	13.1	16.4	21.9	30.9

TABLE 8

GOODNESS-OF-FIT AMONG TUCKER SUBGROUP SOLUTIONS AND WITH
RESPECTIVE FULL GROUP SOLUTIONS (EXAMPLE - SJ12)

	Full Group	Subgrp.1	Subgrp.2	Subgrp.3
Subgrp.1	0.25			
Subgrp.2	0.18	0.39		
Subgrp.3	0.15	0.56	0.49	
Subgrp.4	0.26	0.49	0.49	0.55

TABLE 9

SIGNIFICANT RESULTS OF ANOVA - PERSONALITY
VARIABLES AMONG TUCKER SUBGROUPS

Variable (on association task)	CAU9 F	CAU9 p	CAS9 F	CAS9 p	CAU12 F	CAU12 p
Total Responses	5.2	.009				
Total Internal Keyword Responses	6.7	.001	4.4	.001	18.2	<.001
Fraction Int. Keyword Responses	4.5	.008	3.9	.016	10.4	<.001
Avg. Relatedness Index	16.6	<.001	5.2	.004	12.9	<.001

df error: CAU9 - 43; CAS9 - 38; CAU12 - 32.

TABLE 10

RELATIONSHIPS AMONG CARROLL SOLUTIONS, KRUSKAL SOLUTIONS, AND DATA FOR FULL GROUPS

Data Set	Carroll-Kruskal r	Carroll-Data r
CAU9	0.79	0.60
CAU12	0.75	0.58
CAS9	0.86	0.60
CAS12	0.74	0.58
SJ9	0.95	0.64
SJ12	0.88	0.59
SD9	0.81	0.50
SD12	0.94	0.51

TABLE 11

CARROLL SUBGROUP MEMBERSHIP

Data Set		Subject ID's
CAU9	Group 1	3,10,25,35,38,41,42,61,62,64
	Group 2	2, 9,13,17,27,40,44,54,57,60
	Group 3	2, 6, 8, 9,33,37,43,49,53,63
CAU12	Group 1	11,17,24,27,34,39,49,54
	Group 2	12,17,26,31,40,44,47,48
	Group 3	8,13,14,19,24,29,45,50
CAS9	Group 1	2, 6, 8, 9,12,37,43,49,53,63
	Group 2	2,10,25,35,38,41,42,61,62,64
	Group 3	2, 9,13,17,27,40,44,54,57,60
CAS12	Group 1	11,17,24,27,34,39,49,54
	Group 2	11,12,17,18,26,31,47,48
	Group 3	8,13,14,19,24,29,45,50
SJ9	Group 1	8,9,15,31,33,43,49,57,63,78
	Group 2	2, 3,12,13,25,32,40,44,61,62
	Group 3	21,27,38,39,49,59,65,66,71,76
SJ12	Group 1	4,16,17,21,31,34,49,54
	Group 2	5,19,22,25,29,36,44,45
	Group 3	1, 2, 8,12,15,19,32,44
SD9	Group 1	8,24,26,41,45,55,59,65,66,76
	Group 2	3,12,15,18,27,38,49,67,72,75
	Group 3	7,14,15,17,29,51,61,64,71,74
SD12	Group 1	11,21,22,38,40,41,42,54
	Group 2	6,10,17,28,45,46,47,50
	Group 3	1,15,20,21,26,31,34,52

TABLE 12

GOODNESS-OF-FIT AMONG CARROLL SUBGROUP KRUSKAL SOLUTIONS AND WITH RESPECTIVE FULL GROUP KRUSKAL SOLUTIONS (EXAMPLE - SD9)

	Full Group	Subgrp.1	Subgrp.2
Subgrp.1	0.29		
Subgrp.2	0.24	0.51	
Subgrp.3	0.22	0.41	0.49

TABLE 13

INTERPRETATIONS OF DIMENSIONS IN CARROLL SOLUTIONS OF TUCKER SUBGROUPS

	Grade 9			Grade 12		
Data Set	Dim.1	Dim.2	Dim.3	Dim.1	Dim.2	Dim.3
CAU Group 1	1	2	3	2	4	0
2	2	0	3	1	0	0
3	4	2	3	1	0	0
4	1	4	0	1	0	0
CAS Group 1	1	3	2	0	0	2
2	1	3	0	1	3	0
3	1	3	0	2	0	3
4	4	1	3	4	1	2
SJ Group 1	4	1	3	4	0	0
2	2	3	1	1	3	4
3	1	4	3	4	0	2
4	4	1	2	1	2	4
SD Group 1	1	0	3	0	1	0
2	1	4	0	1	3	3
3	2	4	0	4	3	0
4	1	4	0	1	0	0

0 - no clear interpretation; 1 - certainty; 2 - creativity; 3 - theoretical-practical; 4 - temporal

TABLE 14

FIT BETWEEN KRUSKAL AND CARROLL SOLUTIONS OF TUCKER SUBGROUPS

Data Set	Group 1	Group 2	Group 3	Group 4
		Grade 9		
CAU	0.28	0.08	0.21	0.14
CAS	0.10	0.17	0.19	0.14
SJ	0.16	0.10	0.14	0.27
SD	0.14	0.17	0.11	0.14
		Grade 12		
CAU	0.40	0.32	0.13	0.21
CAS	0.19	0.23	0.23	0.12
SJ	0.20	0.18	0.15	0.18
SD	0.16	0.25	0.26	0.08

TABLE 15

GOODNESS-OF-FIT AMONG CARROLL SOLUTIONS OF TUCKER SUBGROUPS AND CARROLL SUBGROUPS (EXAMPLE - SD12)

	Full Group	Subgroup 1	Subgroup 2	Subgroup 3
		Tucker		
Subgroup 1	0.79			
Subgroup 2	0.77	0.56		
Subgroup 3	0.61	0.50	0.59	
Subgroup 4	0.82	0.62	0.61	0.61
		Carroll		
Subgroup 1	0.76			
Subgroup 2	0.73	0.50		
Subgroup 3	0.79	0.64	0.59	